**U.S. Army Research Institute
for the Behavioral and Social Sciences**

Research Report 1768

The Virtual Sand Table:
Intelligent Tutoring for Field Artillery Training

Robert A. Wisher and Douglas H. Macpherson
U.S. Army Research Institute

L. Jared Abramson and David M. Thornton
George Mason University
Consortium Research Fellows Program

James J. Dees
U.S. Army Training and Doctrine Command

20010405 052

March 2001

U.S. Army Research Institute
for the Behavioral and Social Sciences

A Directorate of the U.S. Total Army Personnel Command

EDGAR M. JOHNSON
Director

Technical review by

Joseph Psotka, U.S. Army Research Institute
Phyllis Robertson, U.S. Army Field Artillery Center and School

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE March 2001	3. REPORT TYPE AND DATES COVERED Interim; May 1998 – May 2000

4. TITLE AND SUBTITLE The Virtual Sand Table: Intelligent Tutoring for Field Artillery	5. FUNDING NUMBERS Program Element: 63007 Project Number: A792 Task Number: 208

6. AUTHOR(S)
Robert A. Wisher and Douglas H. Macpherson (U.S. Army Research Institute),
L. Jared Abramson and David M. Thornton (George Mason University), and
James J. Dees (U.S. Army Training and Doctrine Command)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: TAPC-ARI-II 5001 Eisenhower Avenue Alexandria, VA 22333-5600	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: TAPC-ARI-II 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. SPONSORING/MONITORING AGENCY REPORT NUMBER Research Report 1768

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words):

This report examines the application of an intelligent tutoring system (ITS) for use in training a complex skill during the Field Artillery Captains Career Course. Based on a technology transfer agreement between the U.S. Army Training and Doctrine Command and the U.S. Army Research Institute for the Behavioral and Social Sciences, an ITS originally developed for the Navy was adapted for use in a sand table exercise. The exercise required students to deploy multiple launch rocket system assets during a reconnaissance and selection of position task. The task was conventionally taught using miniature replications of vehicles and launchers on a large table of sand. An ITS version of the exercise, called the Virtual Sand Table, replicated the training with the added advantage of informative feedback and computer-based coaching during the exercise. A comparison group (n=209) used the conventional sand table and the treatment group (n=105) used the Virtual Sand Table during a four-hour training exercise. Results, as measured by a hands-on performance test, indicated superior performance by the Virtual Sand Table treatment group, with an effect size of just over one standard deviation.

14. SUBJECT TERMS training, education, distributed learning, intelligent tutoring, online instruction, field artillery	15. NUMBER OF PAGES 26
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

The Virtual Sand Table:
Intelligent Tutoring for Field Artillery Training

Robert A. Wisher and Douglas H. Macpherson
U.S. Army Research Institute

L. Jared Abramson and David M. Thornton
George Mason University
Consortium Research Fellows Program

James J. Dees
U.S. Army Training and Doctrine Command

**Advanced Training Methods Research Unit
Franklin L. Moses, Chief**
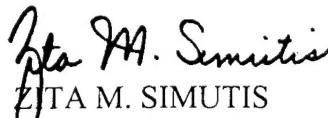
FOREWORD

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is examining the use of distance learning technologies for use by soldiers in an "on demand" environment, where training becomes more soldier centered rather than classroom based. The TRAINTODAY project, sponsored by the U.S. Army Training and Doctrine Command (TRADOC) and now called WEBTRAIN, seeks to provide guidance to the Army as it transforms from a classroom-centric method of instruction to one that is more soldier-centric and more collaborative.

The Field Artillery School had a requirement to transform a segment of complex skill training from the classroom model to an approach that was soldier-centric and could potentially be distributed to remote sites. Matching this requirement, and in accordance with a Memorandum of Agreement between ARI and TRADOC, an intelligent tutoring system applied originally in the Navy was adapted and tested for training effectiveness in the Army. The soldier-centric instructional technology was applied to training complex skills related to reconnaissance, selection, and occupation of position during the Field Artillery Captains Career Course. The research findings reported here were presented to the Assistant Deputy Chief of Staff for Training. TRADOC on 16 March 2000.

ZITA M. SIMUTIS
Technical Director

Acknowledgement

The Virtual Sand Table: Intelligent Tutoring for Field Artillery Training

EXECUTIVE SUMMARY

Research Requirement:

The planned transformation of Army training will lead to a shift from the classroom to training that will become more learner-centric and less instructor dependent. In recognition of this change, the Field Artillery School was interested in transforming the classroom training for the complex skills involved in a sand table exercise to a format that was learner-centric and could be distributed to soldiers through distance learning technologies. The exercise involved reconnaissance, selection, and occupation of position (RSOP) for multiple launch rocket system units.

Procedure:

Through a technology transfer Memorandum of Agreement between the Assistant Deputy Chief of Staff for Training at Headquarters, U.S. Army Training and Doctrine Command (TRADOC) and the U.S. Army Research Institute for the Behavioral and Social Sciences, several off-the-shelf technologies were evaluated for potential transfer and evaluation. A training technology originally developed for radar training in the Navy, ExpertTrain, was selected and a contract awarded to Sonalysts, Inc. to convert it for use in field artillery training. An intelligent tutoring system, called the Virtual Sand Table (VST), was developed for transfer and evaluation in the Army. The VST is learner-centric, allowing training to be conducted at any distance from an instructor. The content of training was the RSOP task, identified by the Field Artillery School as an appropriately important and complex task.

Finding:

An evaluation that compared training performance of those trained on the VST to those trained by a conventional sand table exercise was conducted during the Field Artillery Captains Career Course. End-of-task performance data were collected on n=209 students completing the conventional sand table and n=105 students completing the VST. Results demonstrated superior performance by those trained on the VST, with an effect size of 1.05. On a standard scale, this translates to an improvement for students at the 50th percentile to the 85th percentile of achievement.

Utilization of Findings:

The VST has been implemented at the Field Artillery School, Fort Sill, Oklahoma, for use in the Captains Career Course, and is being used as an instructional aid for the Field Artillery Officers Basic Course. Further improvements to extend the VST training technology to other areas are being pursued by TRADOC.

THE VIRTUAL SAND TABLE: INTELLIGENT TUTORING FOR
FIELD ARTILLERY TRAINING

CONTENTS

List of Figures

The Virtual Sand Table: Intelligent Tutoring for Field Artillery Training

Introduction

Distributed simulations, virtual realities, and intelligent tutors have all reconceptualized distance learning environments during the past decade, reducing the need for an in-place instructor (Dede, 1996). In parallel with these advancements, the U.S. Army Training and Doctrine Command (TRADOC) is embarking on a major change to deliver standardized individual and self-development training to soldiers through the application of multiple media and networked delivery technologies. Training is to move from a classroom-centric delivery of instruction to a learner-centric model, in which soldiers assume greater responsibility for learning facts, procedures, and complex skills in distributed learning environments. The learner-centric model, however, still depends on an instructional source. This may be a qualified instructor available online or an automated tutoring and feedback system.

A training technology that has matured in recent years is intelligent tutoring. As described in more detail later, intelligent tutoring systems have evolved from an arcane art of knowledge engineering and LISP coding to development methods and delivery options that are becoming increasingly mainstream through desktop PCs. Fundamental to an intelligent tutor is a body of domain knowledge encoded as an expert system of rules (Farr and Psotka, 1992). This expertise is accessible to the student during a learning exercise under the control of an instructional strategy. The goal is to have the student construct a mental representation of the domain knowledge -- the expert's facts, rules, and procedures -- for later application.

The purpose of this report is to describe the development and evaluation of an intelligent tutoring system (ITS) for application by the U.S. Army Field Artillery School in the Captains Career Course. Specifically, an ITS was developed for the Multiple Launch Rocket System (MLRS) sand table exercise, a four-hour block of instruction. The training is normally conducted in small groups using a conventional sand table exercise. The Virtual Sand Table is the ITS developed for conducting the same exercise for individuals rather than groups. The present report describes the development of the VST and its evaluation at the Field Artillery School.

This report begins with a review of the origin of the project, followed by a brief overview of the ITS field with examples from military and civilian environments. The development of the Virtual Sand Table is then described. An evaluation of the training effectiveness of the Virtual Sand Table in comparison to the conventional sand table is reported. Finally, plans for application of the Virtual Sand Table in a distance learning format are discussed.

Origin of the Project
The sand table exercise is an intensive training assignment that synthesizes weeks of doctrinal training into a hands-on demonstration of proficiency. It requires the physical presence of an instructor. Transforming this training to a distributed learning format was a challenged recognized by both the Field Artillery School and by the U.S. Army Research Institute (ARI) as part of a research effort to investigate the effectiveness

of "on demand" training environments. Funding for the development of the Virtual Sand Table was an outgrowth of a Technology Transfer Memorandum of Agreement (MOA) between the Assistant Deputy Chief of Staff for Training (ADCST) at Headquarters, TRADOC and ARI. The MOA, which was signed in January of 1997, provides a charter to: (1) investigate new commercial off-the-shelf technologies; (2) assess their applicability to Army training needs; (3) transfer them into Army training; and (4) evaluate their effectiveness. TRADOC was responsible for identifying criteria for selecting technologies for trial in actual Army training programs, while ARI, with its long involvement with emerging training technologies, was charged with identifying high-potential candidates for transfer.

The two organizations collaborated in using a weighted assessment instrument to nominate a training technology developed outside the Army for transfer to the Army. The assessment instrument clustered technology attributes into four categories presented in Table 1.

Table 1. Four training technology rating categories with number of attributes

Rating Category	Number of Attributes
Media and Infrastructure	5
Courseware and Training	5
Logistics and Cost	4
Course Conversion, Implementation, and Management	6

Of the four initial candidates identified for a more detailed assessment, a successful training technology developed by Sonalysts, Inc., ExpertTrain™, scored highest. ExpertTrain is an ITS that creates an intelligent learning environment which is continually modified as a student interacts with a simulation. Its success had been previously demonstrated for the Navy as an intelligent training aid for radar system controllers at the Aegis Training and Readiness Center, Dahlgren, Virginia.

Attributes in each category were assigned a point value based on their criticality to TRADOC's transformation of training to a distributed learning format. For example, under Media and Infrastructure, one of the attributes "Access to Internet" has a maximum point value of six, whereas another attribute "Integration with equipment (may be embedded)" has a maximum point value of three. Integration with actual equipment, while highly desirable, was judged to be less essential for training technologies across the board than was linkage to the Internet. Candidate technologies were rated in all four categories on a total of 20 attributes. ExpertTrain scored the highest.

Origins of Intelligent Tutoring Systems
Early computer instructional systems, termed Computer Assisted Instruction (CAI), aided individual instruction by presenting information in an order that, at best, used simple branching. The path was determined simply by a student's previous answer. Much of the early courseware was passive and oriented to a frame-by-frame presentation of material. Although some of these early systems offered dynamic presentations, they lacked an underlying theory and an instructional approach needed to train skills requiring a higher level of comprehension. Another factor limiting these early systems was the level of computational power available at the time. The early systems were oriented

more towards drill and practice routines. Instruction directed at the development of higher-ordered cognitive skills, such as decision making, analysis, synthesis, and evaluation, had limited capability. Student responses during instruction were generally regarded as independent events rather than as a pattern reflecting a deeper understanding of relationships. As was discovered through intelligent tutors, patterns of responses could also reflect a student's misconceptions about a domain.

An ITS is aimed towards the training of skills and knowledge that are more complex in nature. It is distributed in that it is not limited to the concurrent availability of an instructor. The most important aspect of an ITS is its ability to interact with learners, teaching or assisting them in processing and understanding complex information (Kline,1988). An ITS can generate and customize hints, help, or problems in contrast to the fixed feedback of early CAI systems. An ITS has four main components that distinguish it from other training systems.

> 1. The student module consists of the partial and sometimes incorrect, prior knowledge with which a student begins.
> 2. The expert module contains the correct, expert-level knowledge that is the desired end state.
> 3. The graphical user interface supports two-way interactivity to enable the learning process.
> 4. The instructional module determines what instruction will be given at what point.

The term "intelligent" refers to the system's algorithmic ability to know what to teach, when to teach it and how to teach it. It must have the apparent capacity to understand, and solve problems posed by the student's imperfect or sometimes erroneous comprehension of the domain being taught. An ITS must be capable of identifying a student's strengths and weaknesses and establishing a relevant training path. The instruction can then be tailored to the student's needs in acquiring the expertise as defined in the expert module.

Previous Evaluations of Intelligent Tutoring Systems

The published literature on the effectiveness of distributed learning (DL) is overwhelmingly anecdotal (Wisher & Champagne, 2000). Many evaluations of DL measure only student reactions to the technology, instructor, and course rather than the outcomes of learning. The following section reviews briefly previous research on the effectiveness of intelligent tutoring systems, a special type of DL, that did examine the outcomes of learning. Four representative studies on the effectiveness of ITS are discussed below. The studies concern several domains and can be divided into military and civilian applications.

Intelligent Tutoring Systems in the Military
        Intelligent Tutoring and Device Assembly. A study reported by Orey, Zhao, Fan and Keegan (1998) evaluated the effectiveness of the Intelligently Coached Simulation (ICS) for the assembly of the SINCGARS radio system. An earlier evaluation (Orey, Fan, Park, Tzeng & Gustafson, 1995) showed that an earlier version of the ICS was effective for initial acquisition of the radio-assembly skill, however problems remained

with the transfer (performance on actual equipment) and retention rates (loss of the skill over time). These shortcomings led to enhancements to the ICS.

The enhanced SINCGARS ICS utilized photo-realistic images, an interactive conceptual model, alternative forms of coaching, and tests to overcome the problems of transfer and retention. Its domain knowledge base consisted of the 58 steps to prepare/assemble the system for operation. In both studies, transfer was assessed by measuring the speed of assembly on real equipment. The researchers hypothesized that the addition of photo-realistic imaging of system components would remedy the transfer problem. Retention was assessed by measuring performance on an unannounced post-test four weeks after training. To correct for the earlier retention problem, an improved conceptual model, various forms of coaching and a test during training were implemented.

Orey, et al. (1998) used an experimental design with a post-test only and an additional delayed post-test. The participants were 22 Army officers enrolled in a reclassification course. They were randomly assigned into two groups: the ICS group (n=11) using the SINCGARS tutor, and the hands-on group (n=11) using the actual equipment and a qualified instructor. The post-test consisted of assembling an actual radio for operation. The dependent measures were speed of assembly and number of performance steps done correctly. These were recorded twice: immediately after training and four weeks later as a retention test.

Analyses showed that the ICS tutor was instructionally effective. The computer group performed more accurately than the hands-on group on both the speed and accuracy measures (F= 51.79, p< .001). The skill retention measure compared the performance decline of each group after a four-week retention interval. There was no difference in the relative decline between either group, with an average decline in accuracy of only 4%. The results of this study provide support for the use of a rule-based coach, a conceptual model and a photo-realistic simulation to train students on the operation of equipment.

Expert-based Simulation in the Navy. Van Matre and Robinson (1986) reported on the automated maneuvering board training system (AMBTS), an expert based simulation program designed to train ship handling in the Navy. The AMBTS depicts the geographic location and relative motion of ships to allow training and practice of navigation. A version of the AMBTS was implemented at the Fleet Combat Training Center in San Diego. Operation specialists who had failed the traditional course were provided access to the AMBTS. Previously these students were required to repeat the course with the conventional instruction. However, less than half of these students repeating the class with traditional instruction complete the course successfully. Of those students remediated by the AMBTS (n= 103) 100 percent successfully completed the course. Based on these results, the AMBTS was implemented in the classroom during the skill acquisition portion of maneuvering board training. Instructors requested immediate implementation because of the instructional benefits provided by the system, not its cost-effectiveness.

Intelligent Tutoring in the Air Force. The SHERLOCK is a computer-based coached practice environment employed by the Air Force to train aviation technicians in a realistic context. A study by Lajoie and Lesgold (1992) compared recent cognitive apprenticeship proposals, to coaching via SHERLOCK. The distinction between SHERLOCK and an intelligent tutoring system is that SHERLOCK is not driven by the

student model; its focus instead is on responding to student questions rather than active intervention.

Lajoie and Lesgold (1992) evaluated the SHERLOCK tutor. Participants consisted of trainees (N=63) at two Air Force bases. The experimental design used a two-group pretest-posttest comparison. The control and experimental groups were considered equivalent since the pretest did not yield any differences. Control group subjects worked on their daily activities in the manual avionics shop, while the experimental group spent an average of twenty hours (over 12 days) working with SHERLOCK. Pre- and posttests were conducted in the form of structured interviews.

Analysis revealed that the two groups were significantly different on several competence factors. The experimental group solved significantly more problems than the control group, with respective means of 30 and 21 ($F_{(1, N=62)} = 10.29$, $p < .001$). The tutored group also displayed significantly more expert-like problem solving steps than the control group, with respective means of 19.33 and 9.06 ($F_{(1,27)} = 28.85$, $p < .01$). The SHERLOCK group also made significantly fewer errors ($F_{(1,27)} = 7.54$, $p < .01$). Results demonstrated that subjects who spent twenty to twenty-five hours using SHERLOCK were as competent at troubleshooting as technicians with four more years of job experience.

Intelligent Tutoring in Civilian Settings

Intelligent Tutoring and Abstract Reasoning. Wheeler and Regian (1999) evaluated the effect of the Word Problem Solving (WPS) Intelligent Tutoring System on the abstract reasoning component of word problem solving. The WPS is an adaptive mathematics tutor. Rather than teaching mathematical calculations, it supplements lecture style instruction of pre-algebra, algebra and geometry. The WPS contained sample problems, questions, and summaries of the material. It also varied the type of feedback, difficulty level and number of problems according to the individual skill level of the student.

A sample of ninth-grade students (n=632) was drawn from seven high schools. Students were randomly assigned to one of three groups: a control group (n=84) receiving traditional classroom instruction, a placebo group (n=139) receiving regular instruction, except for one session per week with a non-adaptive tutor session, and a treatment group (n=409) receiving classroom instruction, replacing one session per week with the WPS tutor.

All students were administered a pre- and posttest, designed to measure each student's ability to solve algebra word problems and to compare abstract and concrete reasoning skills. The average scores on the concrete pretests were 29%, 26%, and 33%, for the control, placebo and treatment groups respectively, and 51%, 45%, and 65% for the concrete posttest for the same three groups. The average scores on the abstract pretest were 39%, 33%, and 40% and for the posttest 50%, 49%, and 60%, for the groups respectively. Differences on the gains were shown to be statistically significant at the .001 level. A second test of multiple comparisons indicated the mean change scores for the WPS-tutored (treatment) group on both the concrete and abstract subsets were significantly higher than the mean change scores of the other two groups. This study demonstrated that the WPS tutor significantly improved the performance of high school students on both abstract and concrete word problem-solving subtests.

Development of the Virtual Sand Table

The ADCST, TRADOC selected the MLRS sand table exercise as a model case for evaluating the utility of transforming the Navy-tested training technology for use in the Army. There are many task characteristics shared between the Navy radar training and the MLRS sand table exercise, including the use of higher-order cognitive skills, including analysis, synthesis, decision making, and evaluation. The terminal learning objective of the four-hour sand table exercise is to execute reconnaissance, selection and occupation of position of a firing platoon operations area and battery headquarters using sand tables with associated map sheets. The knowledge base (known also as an

VIRTUAL SAND TABLE ONTOLOGY

Analyze Mission	Receive Orders
	View operations order
Understand Hide Area Criteria	Emplace Hide Area
Emplace Firing Points	Organize Platoon OPAREA
Establish Position	Move To Position
Issue ready reassume control of battery report	Report leave from start point / Start point
Organize Battle Area	

Map reconnaissance	Ground reconnaissance	Select battery movement	Select platoon headquarters pos.
– ID potential positions	– Follow planned route	– Movement routes	– Platoon operations center
– ID potential routes	– Note cover/conceal	– Movement formations	– Ammunition holding area
– ID potential checkpoints	– Note ground conditions	– Movement timing	
– ID potential ambush points	– Consider bridge classification	– Determine release point	
– ID friendly positions	– Check comms		
	– Adjust route as necessary		
	– Notify upon completion		

ontology) developed for the Virtual Sand Table is presented below in Figure 1. The exercise is a critical component of the 18 week Captains Career Course, Phase I, the purpose of which is to prepare field artillery officers for duties as fire support officers at maneuver battalion and brigade level.

Figure 1. Knowledge domain of the Virtual Sand Table

Sand Table Exercise

The conventional sand table exercise is a low fidelity technique used to evaluate the reconnaissance, selection, and occupation of position strategies of soldiers who have completed training on the operation of the MLRS as well as delivery of fires, reload operations, and combat service support. The conventional sand table is an actual table of sand with terrain features molded by hand and MLRS assets depicted by miniature objects. One major limitation of the conventional sand table is its sheer size. There are a limited number of sand tables available at the school. Subsequently, students are compelled to work in groups of five or six, rather than individually as is the case with the VST.

During the sand table exercise, students must review an operations order, evaluate a terrain, and strategically decide where to place firing points, ammunition hiding areas, platoon operations centers, etc. within a terrain model representing up to a

100 square kilometer operations area. The students discuss proposed plans in a group and settle on a consensus plan. After the students have completed setting the sand table with objects that represent positioned assets, one member presents the group's solution to the instructor during an after action review. The plan is then rated by a subject matter expert, and each member of the group is assigned that score, even though some may have contributed little to the plan. The expert judgments are based on the placement of assets and a description of the routes for movement with considerations for the operations order, terrain features, and doctrinal criteria. An example of a doctrinal criterion is that a hiding area must be at least 500 meters from an associated firing point.

The Virtual Sand Table is a computer-based intelligent tutoring system designed to replicate the actual terrain. Terrain maps can be viewed on the screen or fully detailed paper maps can be consulted off-line. Although the conventional sand table may seem a useful training aid, the individual student cannot benefit during the exercise unless there is a highly trained instructor present to critique the process through regular and informative feedback. However, there are few individuals qualified to appraise the performance of these students. When considering the training throughput and instructor resources, an insufficient number of instructors are available to critique students individually during the exercise. This is another reason for grouping students during the conventional sand table exercise. One way to overcome such constraints is for intelligent tutoring systems to provide the critique and feedback throughout the course of the exercise.

Knowledge Engineering

A contract was awarded to Sonalysts, Inc., to develop the VST. As described earlier, the company had previously developed an intelligent tutor for the Navy, using their ExpertTrain intelligent tutoring software. ExpertTrain would be used in the VST. The knowledge engineering phase for the VST is outlined in Table 2.

Table 2. Key Steps in VST Development

Step	Description
1	Obtain all relevant written materials, based on the requirements listed in the statement of work (e.g., the MLRS Field Manual, Firing Platoon Leaders Handbook).
2	Begin an iterative process of discussions with SME's and development of the overall concept of operations, which serve to identify what parts of MLRS activities must be included in the final training product.
3	Once the concept of operations is identified, conduct structured interviews with SME's in order to extract the details of expertise needed to build the coach.
4	Further refine the details of knowledge built into the coach by developing the coaching templates and their triggering conditions, which extensively influenced the simulation design, since the simulation is the source of all student action data.

5	Program the coach (which had to wait until the simulation was under construction) and incorporate access to any quantitative data used in evaluating the student.
6	Review the coach behaviors with USAFAs personnel, tweaking, as necessary, the configurable settings.

Description of Virtual Sand Table

The MLRS VST is essentially a simulation game, where the student's actions are evaluated against a set of expectations that are governed by a set of operational rules. The student receives a mission briefing (OPORD) in the form of text, graphics, intelligence imagery, and a 2-D map display. The software closely replicates standard Army information formats.

The student plans and conducts a map and ground reconnaissance, selection and occupation of position (RSOP). The initial version of the VST is for the duties of a Firing Platoon Leader position. The MLRS VST provides the ability for the student to define points along reconnaissance routes by pointing the mouse on the two-dimensional map display. The terrain defined by the two-dimensional map is pre-processed for each variable in the mission scenario so that the intelligent tutoring component can rapidly assess the student actions. Pre-processing also determines go/no-go/slow-go mobility, major obstructions, bridge classifications, effective hide areas, etc. As the student places assets for occupation of position, the simulation component performs calculations of relative equipment positions and line-of-sight and assesses the student's placement of assets.

The simulation component tracks all of the simulated entities (MLRS launchers, enemy forces, etc.), and their relative positioning with the terrain and each other. The simulation component calculates the line-of-sight, mobility, and trajectories for the MLRS and other vehicles in real-time. The results from the simulation are displayed in the battlefield views and sent to the intelligent tutor component for evaluation.

Two battlefield views are presented to the student: a two-dimensional map which a student may toggle to either a 3km x 3km or a 10km x 10km area view, and a three-dimensional battlefield terrain view. The two-dimensional map view is used to place assets, plan routes, and mark hide areas and firing points. Figure 2 presents the two-dimensional view of one student's partial solution to the problem of emplacing an MLRS Battery in a 3 km area of the National Training Center. Notice that the student has identified 6 firing points (FP 1-6) and a small number of other locations.

The three-dimensional battlefield terrain view is intended for use in the reconnaissance and selection portions of the mission exercise. The student is able to travel through the terrain and observe relevant features and METT-T indicators from the perspective of being in a moving vehicle.

The tutoring component is designed to simulate an instructor coaching a student at a conventional sand table. The focus of the coaching is the evaluation of the student's selected positions and routes in accomplishing RSOP. The basis for the intelligent tutor is a three-step process which includes a situation assessment of the map area 3km grid, diagnosis and evaluation of the student's decisions, and generation of feedback (coaching) to the student. As an elementary example, a student's proposed solution is

provided in Figure 2, and an example of the coaching for that solution is provided in Figure 3.
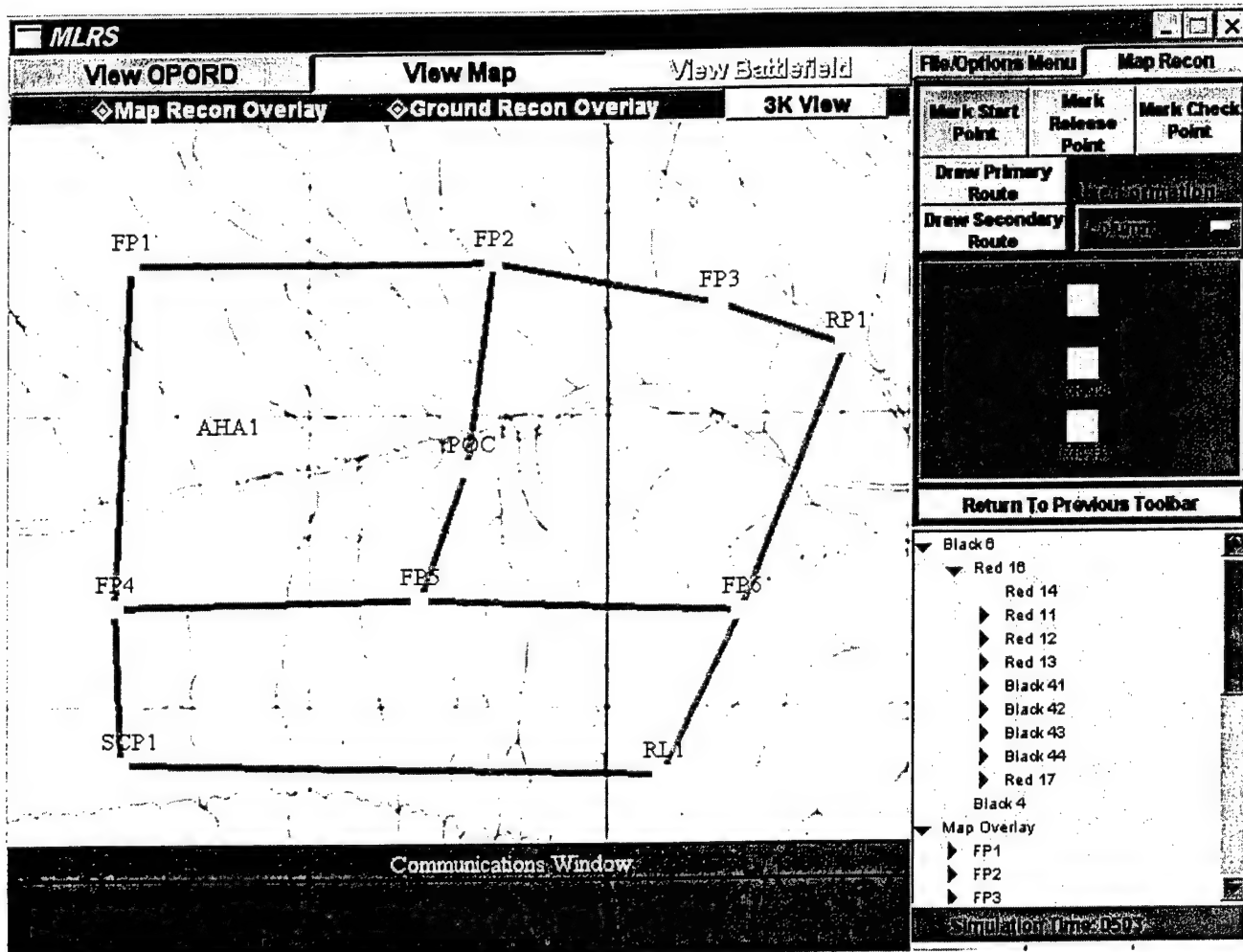

Figure 2. Two-dimensional view with one student's solution to the problem of emplacing a Company in a 3 km area of the National Training Center.
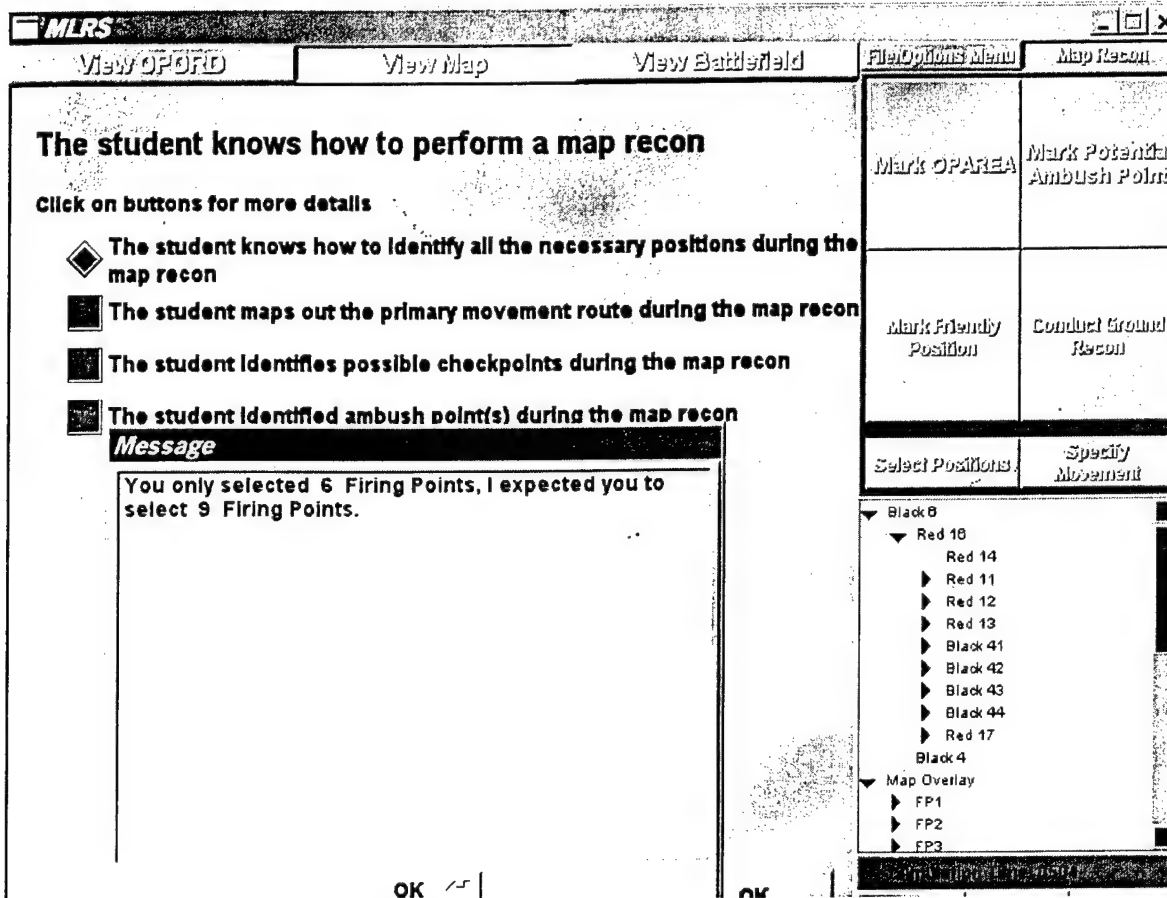
Figure 3. Coaching the student during map reconnaissance; the red diamond icon identifies the subtask causing the message.

The analytical component determines this coaching response by applying rules to the database generated by the student's map operations. The rules are of the form IF FP LESS THAN 9 THEN DISPLAY MESSAGE WITH CONTENTS "You only selected N firing points. I expected you to select 9 firing points." Other aspects of the student's solution examined are:

- 9 Firing Points (FP)
- 9 Hide Areas (HA)
- 2 Reload Points (RP)
- 2 Survey Control Points (SCP)
- 1 Ammunition Holding Area (AHA)
- 1 Platoon Operations Center (POC)
- There is a minimum spacing of 500 meters between firing points and any other point except a hide area

Evaluation of the Virtual Sand Table

The evaluation of the VST was embedded in the normal training schedule for the Captains Career Course (CCC). The conventional sand table exercise is conducted during the fourth week of the course, immediately after students have completed the doctrinal training on the RSOP task. While the VST was undergoing development, five iterations of the CCC students were tracked as a comparison group. The next three iterations of the CCC served as the treatment group. Learning outcomes were measured through a written knowledge test and performance on the sand table exercise.

Participants

Both the comparison group (n=209) and the treatment group (n=105) were given questionnaires before and after the sand table exercise. The first questionnaire sought demographic and experiential information on previous field artillery assignments. The post-exercise questionnaire sought feedback on perceived learning on the same topics and, in the case of the treatment group, ratings on the usability of the VST.

Procedure

In accordance with the Privacy Act, participants were informed of their privacy rights when completing the questionnaires. For the comparison group, soldiers performed the exercise in groups of five. Instructors made a point to include in each group, at least one soldier who had field experience with an MLRS unit. After being provided with an operations order, the group was expected to work together on a solution. Approximately three hours were required by the group to complete the plan. Afterwards, one student was selected at random (through a drawing of straw lengths) to present the plan. The instructor scored performance on a ten-point scale, with ten being outstanding and one being poor. Every member of the group was assigned that score. This score served as the learning outcome measure.

For the treatment group, students were directed to a computer laboratory where the VST exercise was installed on a personal computer. The computers were housed in the Classroom XXI facility at Fort Sill. Each was equipped with a graphics accelerator card. After completing the background questionnaire, an instructor explained the purpose of the VST and then initiated a login procedure. A projection screen displayed the instructor's monitor for viewing by all students. The main features of the VST, such as switching from a top-down view to a 3-D rendition or toggling between a 3 km x 3 km map and a 10 km x 10 km map, were demonstrated. Also described were the graphical-interface procedures for placing MLRS assets and marking routes of travel. The students reviewed an operations order displayed on their monitors. They were given three hours to complete the exercise. Students worked individually.

A primary instructor and three assistant instructors were available to advise students on the mechanical workings of the interface or assist with any technical problems, including an occasional computer crash. Instructors did not advise students on their strategies to select positions or mark travel routes. Upon completion of the exercise, a copy of the plan (a map showing emplacements and routes) was saved electronically in

jpeg format. The same instructor who scored the plans of students in the conventional sand table group applied the same criteria in scoring performance using the same ten-point scale. Scores were recorded individually.

Results

The experimental design for the evaluation was a two-group, multiple post-test design. The results are presented in the following order. First, a comparison of demographic factors between the conventional sand table (comparison) and VST (treatment) groups examines the equivalency between groups, as they were not randomly assigned but rather placed into groups based on the particular period they entered the CCC. Second, the results from a written test that measured the knowledge levels between the two groups are then presented. Finally, scores on the RSOP plans, which measured skill performance, are presented.

Demographic Comparison
The average age of the comparison group was 28.3 years compared to 28.2 years for the treatment group (t=.34, ns). Five other comparisons are presented in Table 3, each demonstrating no significant difference between groups as measured through a chi-square statistic. These results indicate that the two groups were demographically identical.

Table 3. Demographic Comparisons between Groups.

Demographic Variable	Group		χ^2
	Comparison	Treatment	
	Percentages		
Rank			3.41 (not significant)
First Lieutenant	34%	41%	
Captain	66	58	
Major	-	1	
Branch			.36 (ns)
Field Artillery	86	88	
Other	14	12	
MLRS Experience			2.81 (ns)
None	79	76	
Less than 1 yr.	5	2	
1 year or more	16	22	
Cannon Experience			.69 (ns)
None	21	22	
Less than 1 yr.	4	3	
1 year or more	75	75	
Target Acquisition Battery Experience			2.81 (ns)
None	93	94	
Less than 1 yr.	3	3	
1 year or more	3	3	

Knowledge Test

The knowledge test proved to be of no practical value. Individual performance on nine items corresponding to the content of the sand table exercise were extracted from an end-of-module test and analyzed for differences between groups. Both groups scored over 99% correct, reflecting a ceiling effect for that particular measure.

Skill Performance Measure

The comparison group was trained using the conventional sand table approach. The approach consisted of small teams of approximately five students with team assignments that included at least one individual with at least one year of MLRS experience. The students were supposed to work together to solve the problems. Based on feedback from the instructors, the most knowledgeable individual was likely to perform most of the task while those less knowledgeable, or with less rocket experience (or motivation) would hardly participate. This produced the impression that the individuals requiring the least training were obtaining the most from the sand table exercise while those with the least experience and knowledge were least involved. Nevertheless, all students received the same grade.

In contrast, the VST acted as a one-on-one tutor, providing hints and informative feedback to the students individually. To compare the VST students with the groups trained with the conventional approach, it was necessary to create post-hoc groups of students meeting the specifications for composing the conventional training groups. This, in effect, would simulate group performance as derived from individual performance scores. The creation of these post-hoc groups required the application of resampling theory and procedures (Simon, 1995).

Resampling is one of the Monte Carlo based computer intensive statistical techniques. It involves sampling repeatedly, with replacement, from a sample of data. It does not require that the data be drawn from the population at random, but that the sample be representative of the population. The mean of any statistic calculated from the resamples will be the same as that calculated from the data sample. Furthermore, the distribution of the statistic for the resamples themselves will be normal. However, the standard deviation of the statistic will be much narrower. Resampling is different from the usual Monte Carlo approaches in that it employs the original data rather than any summarization of the data.

Of the 105 students who participated in the VST training, 86 were appropriate for analysis. The reasons for exclusion were missing scores for either the MLRS experience variable or the performance score. The program Stats.exe from Resampling Stats was used to create 2,000 virtual groups, each with five students randomly selected. Each group contained at least one student with at least one year of MLRS experience, thus simulating the basis for group assignment in the conventional training procedure. The highest performance score of each group was then assigned to the group. These scores were then compared with the group scores from the traditionally trained groups. The results of these are presented in Figure 4, where a best-fit normalized curve was computed for both groups.
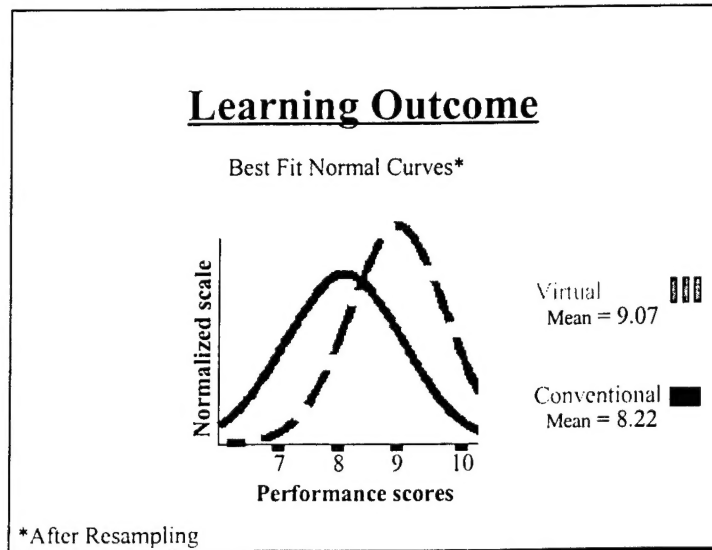
Learning Outcome

Best Fit Normal Curves*

*After Resampling

Figure 4. Normalized curves of performance distribution between groups.

Analysis. The null hypothesis, equal performance between groups, was tested using a two-tailed independent-sample t-test. A two-tailed test was employed because there were no preconceptions about an advantage held by either group. Results from our sample indicated that students trained via the Virtual Sand Table (M= 9.07) significantly outperformed those students trained via the conventional sand table (M= 8.22), $t = 11.43, p < .001$.

Effect Size. The computation of effect size is a common method for determining the gain that a particular treatment has over a comparison group. In this particular case, the effect size is the difference of the mean score for the VST group (9.07) and the conventional sand table group (8.22). When this difference is divided by the standard deviation of the conventional group (as a more valid estimate of the population variance when using resampling), the effect size for the VST is 1.05. This translates to a 35 percent increase in learning (proportion of students surpassing the average performance of the comparison group). The effect size found here is in line with those reported in other studies of intelligent tutors in the military and higher education, which is about 1.0 (Woolf and Regian, 2000).

Discussion

The primary conclusion drawn from this experiment is that the VST tutoring system is an effective tool for training soldiers to perform the RSOP task for the Multiple Launch Rocket System. The results support the VST as a more effective training program than the conventional sand table. Two secondary implications of implementing the VST are reduced training costs and increased accessibility. Before these are discussed, a review of the technology transfer mechanism used to identify and select ExpertTrain for application in an Army training program is presented.

The technology ratings assigned to ExpertTrain at the outset of the technology transfer initiative were assigned prior to its adaptation to Army training. The technology scored well at the outset, but there were several areas that were slightly lower. If the

14

current VST product resulting from this process were to be re-assessed, it would clearly be given higher ratings for "Adaptive instruction," "Instruction based on principles of learning and cognition," "Applicability to the army school system, which includes the reserve component," "Applicability to many soldiers," and "Acceptance of user organization." The tentative rating for "Ease of modification..." would remain the same. The vendor can readily make whatever software changes are required, but because of the proprietary nature of the software, the government pays for modifications.

Although the knowledge engineering phase of the project and the encoding of the expert rules were labor intensive, this intelligent tutoring training program can yield a satisfactory return on investment when training costs are measured longitudinally. In the long run, one of the greatest expenses related to training in the military is travel cost. Traditionally a course or workshop would be offered at a "central" location, and students would travel to the training site. Distributive learning allows students to access information and instruction from their present locations, thus avoiding travel expenses. Since plans are in motion to transform the Captains Career Course to a distance learning format, the VST offers a head start that effectively trains the most complex area of training in the course.

Based on the results reported here, modifications to the VST are underway. The primary modifications consist of enlarging the terrain area and adding capabilities that would allow the role of the Battery Commander (BC) to be expanded. Some of the additional capabilities of the BC would be placing radar units, fire support units, fire direction units, and command post locations on the two-dimensional map. The size of the two-dimensional map would increase from 10 X 10 kilometers to approximately 50 X 50 kilometers for the Battery Commander. The Battery Commander would also be able to map routes from tactical assembly areas to the release point, designate battery operation areas and establish timetables for battery movement.

Distance Learning Plans

The Army has a far reaching plan concerning the establishment of distributed learning sites and the transformation of courses and classrooms to accommodate the distributed training concept. For example, the National Guard Bureau has established high speed network links to armories in all states and territories. Altogether, more than 800 distance training facilities are planned throughout the Army, which would cover 95% of the total force, active and reserve components. In addition to these planned facilities, training will also be delivered to the workplace, to soldiers' residences, or to other sites beyond the traditional classroom. If the performance of students trained via the VST surpasses that of conventionally-trained students, the burdens of travel time and cost, and the problem of limited training slots will be significantly reduced. The Virtual Sand Table can be made available at distance learning sites Army wide, wherever and whenever MLRS RSOP training is needed.

Based on the results of this research, the VST is currently in use as a replacement to the conventional sand table in the Field Artillery CCC and Officer Basic Course. Modifications to the VST have been identified and funding has been provided to further improve its instructional effectiveness and extend its applicability to battery commanders.

Learning Effect. In previous studies on the effectiveness of distance learning media compared to the face-to-face classroom, the "no significant difference" phenomenon (i.e., effect size of zero) has been reported hundreds of times (Russell,

1998). In these studies, the distance learning platform, often video teletraining, replicates the classroom environment. It should come as no surprise, then, that there is not a training advantage for the distance learning media – replicating the classroom replicates the learning outcomes from the classroom. In the present study, the learning conditions of the classroom were not replicated: the training was individualized, feedback during the learning process was customized, the simulation gaming environment was intrinsically motivating, and the training approach was problem based. The large effect size of 1.05 reported here is indicative of the training advantage possible when intelligent tutor technology is applied in a well-designed, learner-centric training environment.

References

Dede, C. (1996). The evolution of distance education: Emerging technologies and distributed learning. *The American Journal of Distance Education, 10*(2), 4-36.

Farr, M. J. & Psotka, J. (Eds.) (1992). *Intelligent instruction by computer: Theory and practice.* Washington: Taylor & Francis.

Kline, K. B. (1988). Intelligent systems for human resources. *Aviation, Space and Environmental Medicine,* November.

Lajoi, S. P., & Lesgold, A. (1992) Apprenticeship Training in the Workplace: Computer-Coached Practice Environment as a New Form of Apprenticeship. *Intelligent Instruction by Computer: Theory and Practice,* pp15-36.

Mark, M. A., & Greer, J. E. (1993). Evaluation methodologies for intelligent tutoring systems. *JL of Artificial Intelligence in Education,* 4(2/3), 129-153.

Orey, M. A., Fan, H., Park, J., Tzeng, S., & Gustafson, K. (1995). Evaluation of device operation in the context of a coached simulation environment. (Contract No. DAAL03-91-C-0034, TCN Number 94-021, Delivery Order No. 1104). Washington, DC: United States Army Research Institute.

Orey, M., Zhao, R., Fan, H., & Keenan, R. (1998). Summative evaluation of the SINCGARS tutor. *Computers in Human Behavior,* Vol. 14, No. 4, 579-595.

Russell, T. L. (1998). The "no significant difference phenomenon. [Online] Internet: http://tenb.mta.ca/pheno/phenom.html

Simon, J. L. (1995). *Resampling: The new statistics.* Arlington, VA: Resampling Stats, Inc.

Van Matre, N., & Robinson. (1986). The automated maneuvering board training system: An inprogress implementation evaluation. *Journal of Of Computer-Based Instruction,* Vol. 13. 2, 35-38 .

Wheeler, J. L., & Regian, J. W. (1999). The use of a cognitive tutoring system in the improvement of the abstract reasoning component of word problem solving. *Computers in Human Behavior,* 15, 243-254.

Wisher, R.A. & Champagne, M.V. (2000). Distance learning: An evaluation perspective. In S. Tobias and J. Fletcher (Eds.), *Training & retraining: A handbook for business, industry, government, and the military* (pp. 385-409). New York: Macmillan Reference USA

Woolf, B.P. & Regian, J.W. (2000). Knowledge-based training systems and the engineering of instruction. In S. Tobias and J. Fletcher (Eds.), *Training & retraining: A handbook for business, industry, government, and the military* (pp. 339-356). New York: Macmillan Reference USA